

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re patent application of: ASHIDA, et al.

Serial No.: Unknown

Group No.: Unknown

Filed: Unknown

Examiner: Unknown

For: **METHOD AND SYSTEM FOR DATABASE MANAGEMENT FOR DATA MINING**

Assistant Commissioner for Patents
Washington, D.C. 20231

PRELIMINARY AMENDMENT

Sir:

Please make the following change prior to examination of the above-referenced application:

In the Specification:

Please insert following sentence between “THE TITLE OF INVENTION” and “Field of the Invention:”

- -This is a continuation of prior application Serial No. 09/994,951 filed on November 27, 2001 under 35 C.F.R. 1.53(b).- -

Please amend the following:

Page 8, lines 12-13, between “the rule,” and “A precision”, please insert --200 people satisfy the rule portion while 50 people satisfy both the rule and condition portions.--

Page 10, line 15, between “108.” and “Next,” please insert - -The month for the above analysis is automatically selected to be two months after the currently selected month. As described above with respect to FIGURE 2, certain portions of the data other

**METHOD AND SYSTEM FOR DATABASE MANAGEMENT FOR DATA
MINING**

This is a continuation of prior application Serial No. 09/994,951 filed on November 27,

5 2001 under 35 C.F.R. 1.53(b).

Field of the Invention

The current invention is generally related to a database analysis technology, and
10 more particularly related to the generation of a customer list based upon a certain
predetermined purpose using a speculation model.

BACKGROUND OF THE INVENTION

15 In the recent years, magnetic cards and IC cards have been widely used in
combination with computer equipment. With the above cards, customer databases have
been developed and maintained in various industries such as department stores, specialty
boutiques, consumer electronics retailers and super markets. The above databases include
customer characteristic information such as names and addresses as well as other
20 information such as accumulated purchase data. Similarly, transactions are maintained in
the databases for the financial industry while data called call detail data are maintained in
the databases for the telecommunication industry. For example, the call detail data include
a caller number, a recipient number and call duration for each call. Based upon the above
described databases, one exemplary service is Customer Relationship Management (CRM)
25 for providing quality service.

Another exemplary use of the above described databases is data mining that semi-
automatically extracts certain information by analyzing a large volume of database data. In
particular, data mining includes rule induction, Memory Based Reasoning (MBR), On-
30 Line Analytical Processing (OLAP), and the these exemplary data mining methods are
disclosed in "Data Mining Techniques For Marketing, Sales and Customer Support," pp.
120-123, John Wiley & Sons, Inc (1997). Rule induction generally extracts certain

classifications of the layer structure defines a speculated segment that is a portion of data to be speculated. In the above example, the speculated segment is a portion of the customer data that is defined by the above described combined conditions. For example, the speculated segment is expressed by age = 20~24 & gender = female & profit amount = \$300~\$400. One rule generation technique is disclosed in "Proceedings of 1999 IEEE International Conference on Systems, Man, and Cybernetics," p.V.-882-886.

Now referring to FIGURE 4, one example of the characteristic rule sets 104 is illustrated based upon the customer 101 in the March data. A first column includes entry items such as numbers while the rest of the columns each includes one rule. A rule sentence in the second column is written in the "if ... then" format. For example, if the age is between twenty and twenty-four and the gender is female, license is cancelled. A rule/condition in the third column is a ration between a number of records to satisfy the rule and a number of records to satisfy only the condition portion of the rule. 200 people satisfy the rule portion while 50 people satisfy both the rule and condition portions. A precision level in the fourth column is a ratio between the number of satisfying records for the rule and the number of satisfying records for the condition.

Now referring to FIGURE 5, an exemplary multidimensional display is illustrated. In this example, the above rule No. 1 is selected in FIGURE 4. The selected rule is that if the age is between twenty and twenty-four and the gender is female, license is cancelled. Based upon the above selected rule, a multidimensional display screen displays condition items as well as conclusion items, and the multidimensional display includes rows for displaying age groups and columns for displaying gender. In each cell, the above ratio between the number of cancelled customers for the rule and a total number of customers is displayed as a conclusion item. The above ratio value is automatically calculated by the system according to the current invention. The cells that meet the conditions used in the selected rule are in a certain predetermined color in order to distinguish at a first glance from other conditions that are not used in the rule. Other conditions are displayed as pages of the multidimensional database.

segment is worthwhile for predictions. An example of deleting a condition in the above example to restore the profit amount to the originally selected all-amount condition. As described above, the user focuses upon a certain cell after he or she adds or deletes conditions to see the cell values in the certain cells and cells around the certain cells.

5

Still referring to FIGURE 6, after the user added the condition on the profit amount of \$300-\$400 in combination with the existing conditions of age = 20 through 24 and the gender = female, the above conditions determine the selected segment 108 as shown in FIGURE 1. Using a pointing device such as a mouse, a particular cell is selected as a target cell for speculation. Furthermore, a set of predetermined functions is also displayed for the selected cell when the user initiates the menu. For example, the menu display is initiated by a right mouse button while the cell is selected by a left mouse button.

Within the function menu, the user selects a desired function by the left mouse button. Assuming that the user selects the selected customer list generation in the function menu and the March data is currently being displayed, the selected customer list 107 is selected from the customer data 101 from May or two months after the current data and only from a portion that satisfies the imposed conditions 108. The month for the above analysis is automatically selected to be two months after the currently selected month. As described above with respect to FIGURE 2, certain portions of the data other than a specified data such as the cancellation status are automatically taken from two month earlier. Next, assuming that the user selects the speculation mode generation in the function menu, the speculation model generation unit 109 automatically generates an optimal speculation model based upon the conditions that the user has selected for the above described segment selection process or unit 106. Lastly, assuming that the user selects the speculation in the function menu, the speculation processing unit 111 automatically ~~conclude~~ concludes the speculation results 112 based upon the selected customer list 107 and the speculation models 110. The speculation algorithm is substantially the same as the algorithm used for speculating the potential cancelled customers or possibility for the cancelled customers. The speculation algorithms include prior art techniques that have been disclosed in the background section of the current application. The speculation item in the function menu remains disabled until the selected customer list 107 and the speculation models 110 have been selected and successfully completed.